Indian J. Phys. **51B**, 462-472 (1977)

Vibrational spectra of oligomers of glycine, alanine and proline in relation to their polymeric forms

M. K. GUPTA AND V. D. GUPTA

Physics Department, Lucknow University, Lucknow, U.P.

The frequency versus phase relationship for a finite chain with free ends and the extended phonon dispersin curves for poly glycine, poly-L-alanine and poly-L-proline in different conformations have been used for the vibrational analysis of oligomers of glycine, alanine and proline respectively. It is observed that the oligo glycines upto the pentamer and oligoalanines upto the hexamer are in a linear zig conformation whereas those of proline after the tetramer have a conformation nearer to the right handed three fold helical structure like poly-L-proline II. Hexaglycine and the higher oligomers also tend to go into a helical structure nearer to that of polyglycine II. These observations are also supported by a progressive change in IR and Raman active conformational sensitive modes and other studies.

1. INTRODUCTION

The conformation of macromolecules and polypeptides has been a subject of biological as well as theoretical interest and several conformation sensitive techniques have been used to study it[1]. Synthetic homooligopeptides provide good, simple models for the study of more complex proteins. They have an added advantage over model compound because they include the effects of both the chains length and the nature of side chains on the stability of polymeric conformation. Such studies are also helpful in mapping the dispersion curves of one-dimensional chains[2] to interpret the spectra of short chain molecules having the same conformation as the polymeric forms[3,4,5]. In the present communication we use the extended phonon dispersion curves for the polymeric forms to study the vibrational spectra of oligopeptides of glycine, L-alanine and L-proline.

Glycine is the simplest anino acid and it's polymeric forms i.e. poly glycine exists in two crystalline modifications designated as I & II. Form I is a planar zig-zag with a two fold screw axis symmetry, whereas Form II has a helical

structure with a three fold screw axis. Poly-L-Alanine is the simplest polypeptide which exists both in the alpha helical as well as beta (planar zig-zag) conformation. The alpha form is a right handed helical structure with a turn ratio of 18/5. The helix is stabilized by intrachain hydrogen bonding and crystallizes in a hexagonal arrangement.[6] The beta form contains fully extended linear zig-zag chains with a twofold screw axis like -(Gly)-$_n$ I. A progressive change in the intensity and positions of IR and Raman bands have been observed by Shotts and Sievers,[7] Simon *et al.*[8] and Sutton and Koenig.[9] Poly-L-Proline also exhibits two forms I & II. The form I assumes a right handed helical structure and has each peptide bond in a cis conformation with an axial translation of 1.90 Å.[10] In the form II peptide bonds are in a trans configuration. The trans helix is left handed with three residues per turn and an axial translation of 3.12 Å per residue.[11]

2. Theory

The infrared spectra of short chain molecules having identical units are usually characterized by a series of absorption bands called as band progression. The normal mode responsible for a component absorption in the band progression can be specified by the phase difference among the adjacent units. Fanconi *et al*[27] have given a theoretical treatment for it; however a brief review is given here. For helical polymers the Wilson's G and F[28] matrices as modified by Higgs[29] are

$$G(\Phi) = G^{\circ} + \Sigma e^{is\phi} G^{s} + e^{is\phi} \tilde{G}^{s} \quad \ldots\ldots 1(a)$$

$$F(\Phi) = F^{\circ} + \sum_{s}^{s} e^{is\phi} F^{s} + e^{-is\phi} \tilde{F}^{s} \quad \ldots\ldots 1(b)$$

φ is the phase relationship between vibration in adjacent unit cells, the tilde indicates the transpose of the matrix. If there are 3N cartesian coordinaestes in a unit cell, the 3N solutions of the secular equation

$$| G(\varphi) F(\varphi) - W^2 I | = 0 \quad \ldots\ldots 2)$$

for each value of φ, give the phonon dispersion curves.

$$w^2 = 4 (f/m) \sin^2 \varphi/2 \quad \ldots\ldots (3)$$

where $0 \leq \Phi \leq \pi$

On the other hand for a linear chain of N identical particles of mass m, connected by springs of force constants f, the eigen value equation has (N-1) solutions

$$w^2 = 4\ (f/m)\ \sin^2\ (L\pi/2N) \quad L = 1,2, \ldots N\text{-}1. \qquad (4)$$

On comparing equations 3 & 4

$$\varphi = L\pi/N \qquad \ldots\ldots\ 4(a)$$

for a finite chain with free ends. However if the ends are rigidly fixed

$$\varphi = \frac{L\pi}{N+1} \qquad \ldots\ldots\ 4(b)$$

where L represents the number of loops in a stationary wave representing the chain vibration. In case of weak hydrogen bonds at the chain ends, equation (4a) can be used with good approximation.[30] It is assumed valid in oligo-systems reported in this communication.

3. Results and Discussion

(a) *Oligoglycines* :

A complete normal mode analysis and the disperson relations of -(Gly)-$_n$ I & II have been reported by several workers[12-15]. The Raman scattering data on oligoglycines upto the pentamer, as reported by Smith *et al.*[16] has been used. The infrared spectra of hexa and dodecamer have already been reported by Dwivedi *et al*[17]. A comparison of corresponding modes in various oligomers and the two forms of -(Gly)-$_n$ has been given in Table 1. It is evident from the table that the oligomers upto hexamer have frequencies nearer to the form I whereas dodecaglycine has corresponding frequencies very close to form II. As an example the skeletal stretch frequency occurs around 1010 cm^{-1} for $n \leq 6$ and at 1016 cm^{-1} in the form I, whereas in dodecaglycine and form II the same mode occurs at 1029 cm^{-1} and 1028 cm^{-1} respectively. Not only amide or skeletal frequencies, methylene group frequencies for oligoglycines upto hexamer are also nearer to the form I, and for dodecaglycine are closer to the form II, e.g. CH_2 bending and wagging modes for hexa and dodecaglycine appear at (1432 and 1408) and (1419 and 1377 cm^{-1}) respectively which are close to those found in -(Gly)-$_n$ I and II respectively. A detailed comparison of amide, skeletal and methylene group frequencies and a good fit of observed frequencies for oligoglycines on the corresponding dispersion curves for allowed phase values show that the oligomers of glycine upto hexamer have a structure closer to the extended beta form whereas dodecaglycine has a right handed three fold helical structure. Since the oligoglycines from hepta to undeca are not available, it is difficult to say at what chain length the oligomeric form goes from type I to II structure.

(*b*) *Oligoalanines* :

A complete normal mode analysis and the dispersion relations for alpha and beta poly-L-Alanines have been reported by Itoch and Shimanouchi[18] and by Gupta *et al.*[19] respectively. The corresponding modes in various oligomers

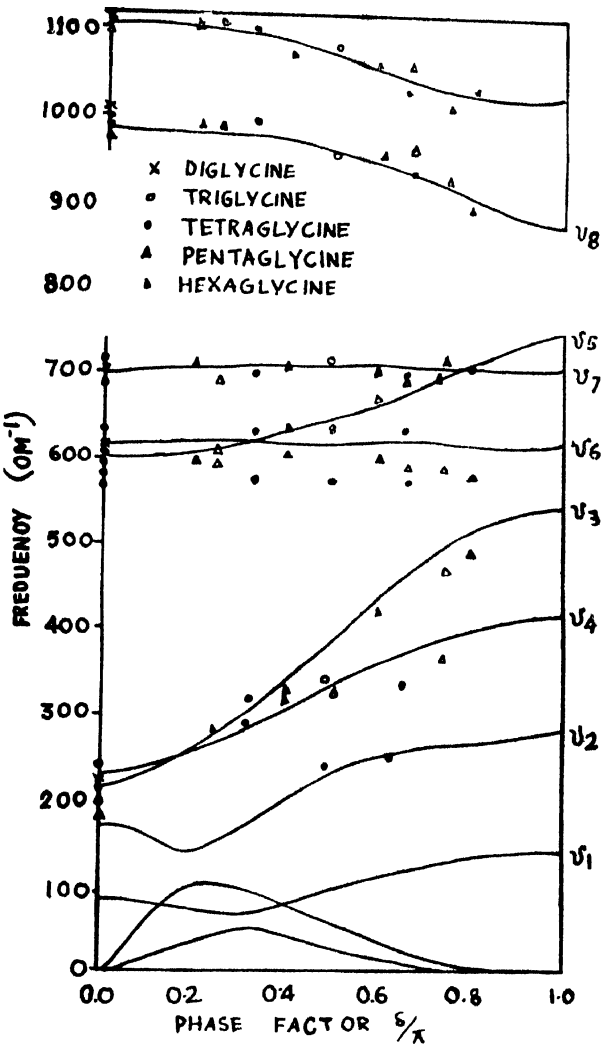

Figure 1. Dispersion curves for PGI.

and the two forms of poly-L-alanine are compared in Table 2. Most of the characteristic frequencies of the oligomers fall in the neighbourhood of those for the beta form. The agreement becomes progressively better for higher number of residues in the chain. There is very little divergence for $n > 4$.

This shows that although the minimum number of residues for the formation of a turn of the helix is present, the oligo alanines still tends to go (in the solid state) into a structure nearer to the beta form. This is unlike the oligoglycines, which go into a helical conformation.[20] In case of di-alanine there are signifi-

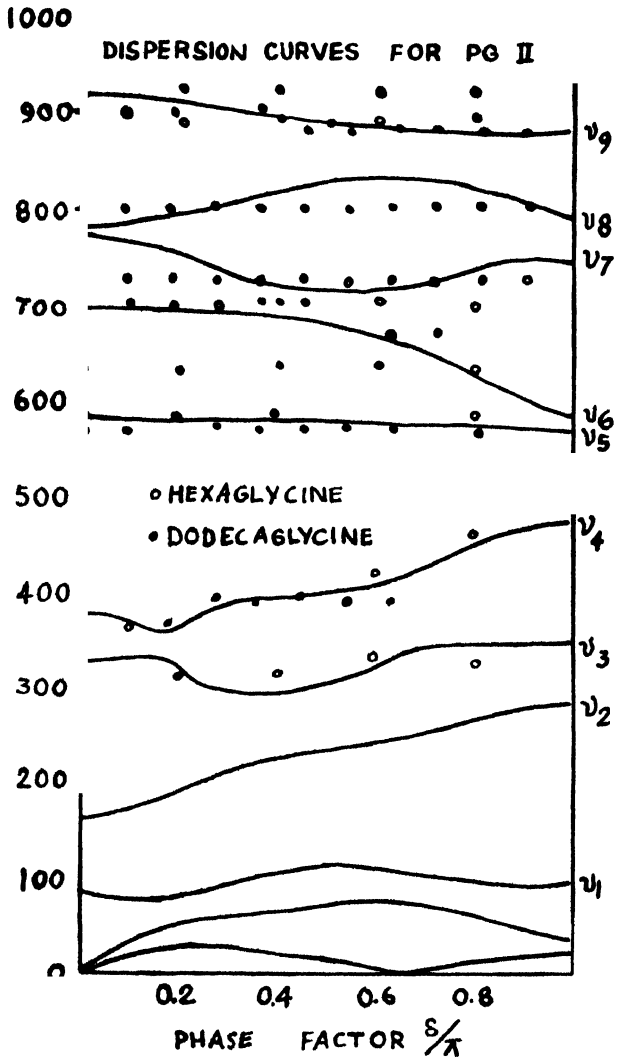

Figure 2. Dispersion curves for PGII.

cant deviations. These could be due to either the effect of the end groups or the out-of-plane conformation of the COO group with respect to the plane of the atoms in the skeleton (N—C_α—C—N—C). The importance of the end effects progressively goes down for higher n values. It may be mentioned here that in solution of chloroform—trifluoroacetic acid (TFA), the nmr

spectra[21,22] of a series of alanine oligomers (pentamer to nonamer) exhibit "double peaks" for the α-CH protons which on further addition of TFA coalesces into a single peak. This shows that in solution the oligopeptides go into a specific folded form stabilised by intramolecular hydrogen bonds. It may be nearer to alpha (18/5) or gamma (36/7) forms. In solid phase the situation is, however, different. The X-ray photographs show the presence of a diffraction peak at 5.33 Å which increases in intensity with increasing n and is typical for poly-alanine in β-form.

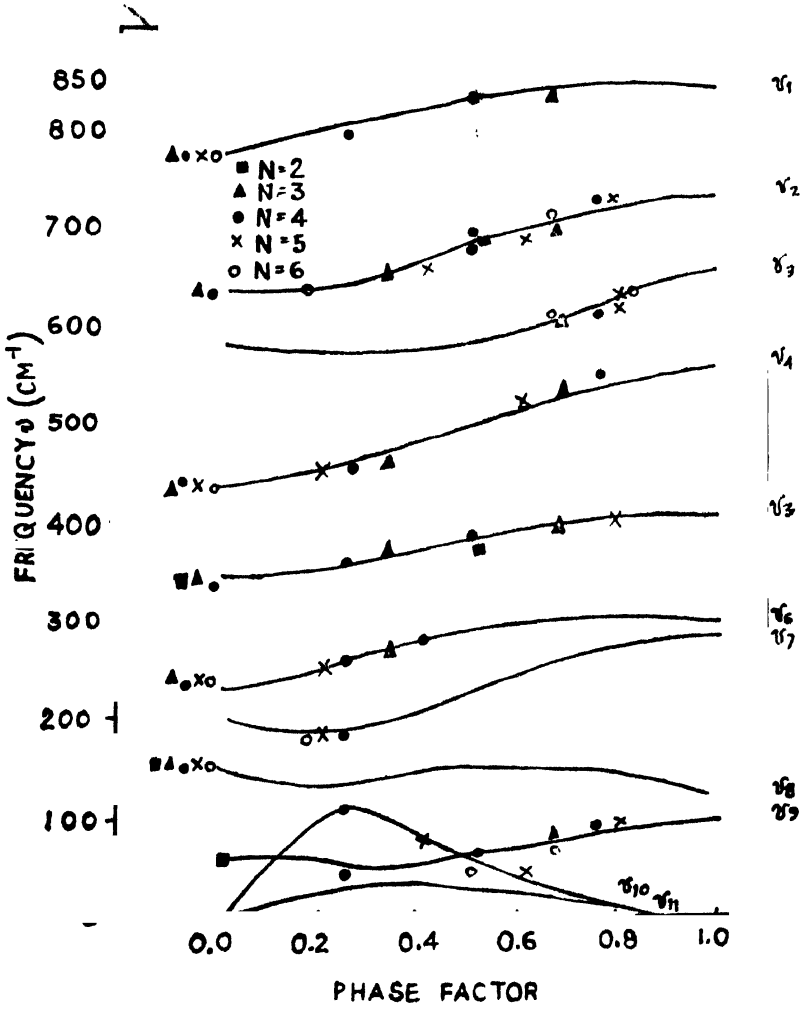

Figure 3. Dispersion curves for β poly-L-aniline.

The vibrational spectra and dispersion curves of Poly-L-Proline in forms I and II have already been reported[23,24]. The corresponding modes in various oligomers and the two forms of poly-L-proline are compared in Table 3. The

observed frequencies for the oligomers again correspond very nearly to K value on the dispersion curves. The agreement is better for the greater number of residues in the chain e.g., the skeletal bend frequency which appears as a very weak band around 312 cm^{-1} in pentamer becomes a broad and intense

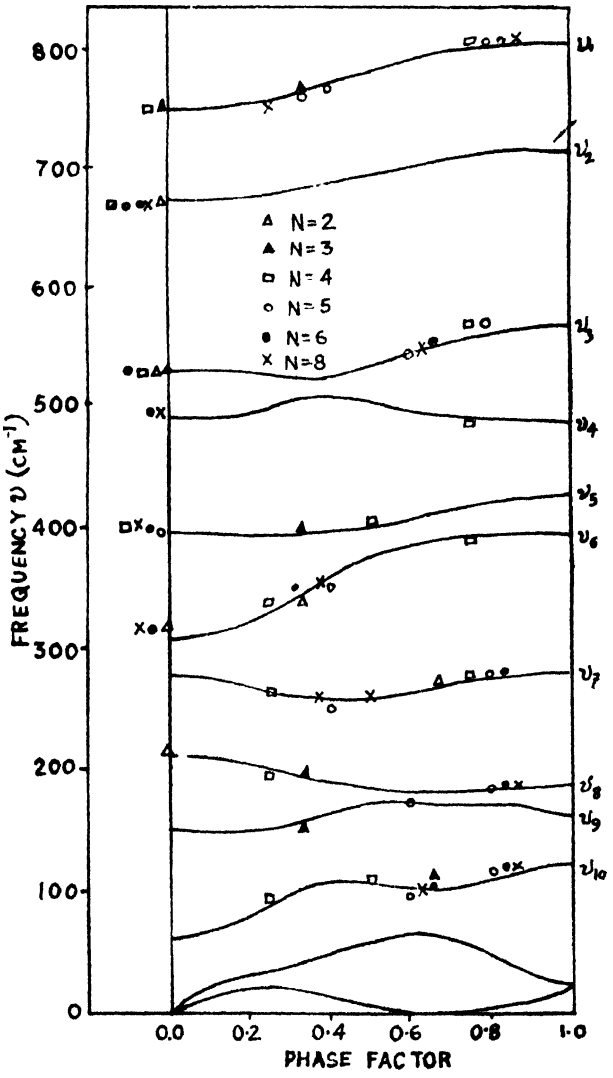

Figure 4. Dispersion curves for PLPII chain.

band in the octamer and is very close to the corresponding band in poly-L-proline II. Thus from a study of the dispersion curves it appears that $(L\text{-}Pro)_4$- goes into a trans helical form (minimum number of residues needed for the for formation of a turn of the helix) and at the octamer stage, the

Table 1. Amide, skeletal and methylene group frequencies in oligomeric glycines and polyglycine

Modes	n=2	n=3	n=4	n=5	n=6	n=12	Poly Glycine	
							Form I	Form II
Amide I	1645	1675, 1635	1685, 1635	1683, 1630	1682, 1632	1650	1685, 1636	1654
Amide II	1548	1550, 1515	1522	1520	1525, 1545	1555	1534	1554
Amide III	1243 1233	1282, 1245	1240, 1220, 1215	1243, 1230, 1214	1225, 1212	1250, 1202	1236	1283, 1244
Amide IV	615(600)	645(620)	640(645)	633(628)	635(620)	700(698)	628	698
Amide V	708(700)	718(706)	700(700)	712,(706) 695	705(705)	745,(746) 724(725)	708	740
Amide VI	588(615)	585(620)	575(610)	585(605)	596,(600) 582	568(570)	589	573
Amide VII	—	—	—	—	—	365	217	363
CH_2 Bend	1445	1429	1435	1432	1432	1419	1432	1420
CH_2 Wag	1408	1395	1402	1403	1400	1375	1408	1384
Skeletal Stretch	1007(992)	1006(992)	1006(992)	1011(992)	1012(992)	1029(992)	1016	1028

All frequencies are in cm^{-1} and the figures in parenthesis are the predicted values from the dispersion curves of form I for n=2, 3, 4, 5, 6 and from the dispersion curves of form II for n=12.

Table 2. A comparison of amide group and skeletal frequencies in $(L\text{-ala})_n$ ($n=2, 3, 4, 5, 6, \infty$)

Modes	n=2	n=3	n=4	n=5	n=6	Poly-L-Alanine	
						α-form	β-form
		1692	1696	1698	1696		
Amide I	1685	1666	1679	1670	1650	1659	1695
		1639	1662	1652	1631		
			1633	1632			
		1542					
Amide II	1539	1536	1539	1539	1538	1515	1524
		1232	1248	1243	1237		
Amide III	1233	1219	1223	1221	1221	1270	1224
Amide IV	651(657)	631(631)	628(631)	625(631)	624(635)	655	629( )
Amide V	733	694(699)	690(685)	692(700)	701(699)	756	705( )
Amide VI	554	606(612)	609(621)	615(632)	610(605)	595	594
Amide VII	240(238)	243(238)	240(238)	241(238)	241(238)	185	247
Skeletal stretch	849	771(773)	772(773)	768(773)	770(773)	748	773
Skeletal bend	—	528(535)	558(561)	521(514)	—	367	—
	—	434(437)	439(437)	436(437)	435(437)	—	437
	404	399(390)	—	406(408)	—	—	—
	365(380)	371(365)	380(380)	—	—	—	—
	339(349)	345(349)	340(349)	—	—	286	349
Skeletal torsion	151(151)	154(151)	150(151)	155(151)	155(151)	163	151
	97	92(84)	86(83)	92(85)	—	120	88
	75(65)	79(76)	68(65)	—	74(76)	113	65

All frequencies are in cm^{-1} and the figures in parenthesis are the predicted values from the dispersion curves of Poly-L-Alanine in β-form.

Table 3. A comparison of ring modes and skeletal frequencies in $(L\text{-}Pro)_n$ ($n=2, 3, 4, 5, 6, 8, \infty$)

Modes	n=2	n=3	n=4	n=5	n=6	n=8	Poly-L-Proline	
							Form I	Form II
Skeletal stretch	795(747)	805(763)	805(798)	805(800)	805(802)	805(804)	960	820
Ring bend	685(670)	677(670)	675(670)	670(670)	670(670)	670(670)	614	670
Ring torsion	545(528)	527(524)	538(525)	545(540)	540(545)	550(542)	540	535
	510	—	510	—	508	505	—	495
Skeletal bend	—	410(400)	405(400)	400(400)	405(400)	405(400)	—	400
	—	—	375(380)	350(345)	350(345)	350(360)	363(R)*	—
	—	—	350	312	315(314)	315(314)	280	314
	294	290	—	282(276)	276(278)	285	—	285
Skeletal torsion	268	—	262(264)	258(262)	262(262)	—	—	260
	213(211)	203(200)	197(202)	190(190)	190(191)	187(191)	160	190
	—	100(100)	98(103)	100(102)	99(102)	97(100)	112	100

* R stands for Raman frequencies.

All frequencies are in cm^{-1} and the figures in parenthesis are the predicted values from the dispersion curves of Poly-L-Proline II

helix is well stabilised and its characteristic modes are very nearly those of poly-L-proline II. This is supported by recent conformational calculations of Hopfinger and Walton[25]. Lastly, it may be added that the laser excited Raman spectrum of poly-L-proline II in the solid state is very similar to that of poly-proline in aqueous solution in which form II is stable[26].

References

1. Walton A. G. & Blackwell J. *Biopolymers*, Academic Press, New York, 1973.
2. Snyder R. C. & Schactschneider J. H. *Spectrochim. Acta*, **19** (1963), 85.
4. Uno T. & Machida K. *Spectrochim. Acta* **24** (1968) 1741.
5. Greine Y. (1968) *Ph. D. thesis submitted to La Faculi'e des Sciences, Universite dc Bordeus* (France).
6. Elliot A. & Malcolm B. R., *Proc. Roy. Soc.*, **A249** (1959), 30
7. Shotts W. J. & Sievers A. J. *Biopolymers* **13** (1974) 2593.
8. Simons L., Bergstrom. G., Blomfel G., Fross S., Stenback H and Wansen G., *Commentatione Physico—Mathematicae*, **42(3)** (1972), 125.
9. Sutton P. L. & Koenig J. L. *Biopolymers* **9** (1970) 615.
10. Traub W. & Shmueli W. W. in *Aspects of Protein Structure*, G. N. Ramachandran, Ed., Academic Press, New York, 1963 p. 81.
11 Sasisekharan V. *Acta Cryst.*, **12** (1959), 897.
12. Gupta V. D., Trevino S. and Boutin H. *J. Chem. Phys.* **48** (1968), 3008.
13. Small E. W., Fanconi B. & Peticolas W. L. *J. Chem. Phys.* **52** (1970), 4369.
14. Singh R. D & Gupta V. D *Spectrochim. Acta* **A27** (1971) 385.
15. Abe Yasuaki and Krimm S *Biopolymers*, **11** (1972) 1841.
16. Smith M., Walton A. G. & Koenig J. L. *Biopolymers*, **8** (1969), 29.
17. Dwivedi A. M. & Gupta V. D. *Chem. Phys. Letters* **8** (1971) 320.
18. Miyazawa T., Fukushima K. Sugano S. and Masuda Y. in Conformation of *Biopolymers* Vol. 2. G. N. Ramachandra, Ed., Academic Press, New York, 1967, p. 557.
19. Gupta, V. D. & Krishnan M. V. *J Phys.*, **B3** (1970), 572.
20. Gupta V. D., Gupta M. K and Nath K. *Biopolymers*, **14** (1975), 1987.
21. Goodman M., Ueyama N. & Naider F., *Biopolymers* **14** (1975), 901.
22. Goodman M., Ueyama N., Naider F & Gilon C. *Biopolymers* **14** (1975), 915.
23. Dwivedi A. M. & Gupta V. D. *Chem. Phys. Letters* **16** (1927), 109.
24. Gupta V. D., Singh R. D. & Dwivedi A. M. *Biopolymers* **12** (1973), 1377.
25. Hopfinger A. J. & Walton A G. *J. Macromol. Sci. Phys* **3(1)** (1969), 171.
26. Smith M., Walton A. G. & Koenig J. L. *Biopolymers* **8** (1969) 173.
27. Fanconi B., Small E. W. & Peticolas, W. L. *Biopolymers* **10** (1971), 1277.
28. Wilson (Jr) E. B., Decius J. C. & Cross P. C. *Molecular Vibrations* (McGraw-Hill Book Co., Inc., New York), 1955.
29. Higgs P. W., *Proc. Roy. Soc. London*, **A220** (1953), 472.
30. Wu C. K. & Nicol M. *J. Chem. Phys* **58** (1973) 5150.